

## THETA METHOD DYNAMICS

GRAEME J. BARCLAY, DAVID F. GRIFFITHS AND DESMOND J. HIGHAM

*Abstract*

Long-term solutions of the theta method applied to scalar nonlinear differential equations are studied in this paper. In the case where the equation has a stable steady state, lower bounds on the basin of non-oscillatory, monotonic attraction for the theta method are derived. Spurious period two solutions are then analysed. Under mild assumptions, precise results are obtained concerning the generic nature and stability of these solutions for small timesteps. Particular problem classes are studied, and direct connections are made between the existence and stability of period two solutions and the dynamics of the theta method. The analysis is extended to a wide class of semi-discretized partial differential equations. Numerical examples are given.

## 1. Introduction

When applied to a scalar autonomous ordinary differential equation (ODE)

$$u' = g(u), \quad u(0) = u_0, \quad (1)$$

the theta method takes the form

$$u_j = u_{j-1} + \Delta t(1 - \theta)g(u_{j-1}) + \Delta t\theta g(u_j). \quad (2)$$

Here  $u_j$  is the numerical approximation to  $u(j\Delta t)$  and  $\Delta t > 0$  is a constant timestep. We assume that the fixed parameter  $\theta$  is chosen so that  $0 \leq \theta \leq 1$ . For  $\theta \neq 0$  the formula (2) is implicit in the unknown  $u_j$ , and hence, in general, a nonlinear equation solver must be used at each step. On a constant coefficient linear problem, where  $g(u) = \lambda u$  in equation (1), we have

$$u_j = R(\lambda\Delta t)^j u_0,$$

where

$$R(z) = \frac{1 + (1 - \theta)z}{1 - \theta z} \quad (3)$$

is known as the *stability function* of the method.

Note that  $\theta = 0$  in equation (2) gives Euler's method,  $\theta = \frac{1}{2}$  gives the trapezoidal rule and  $\theta = 1$  gives the implicit or backward Euler method; see, for example, [3]. Each of these methods is widely used in the context of solving initial value ODEs and, more generally, for timestepping in the solution of partial differential equations (PDEs). The trapezoidal rule is a second-order method, whereas for  $\theta \neq \frac{1}{2}$  first order is achieved. In some applications,

---

All three authors were supported by the Engineering and Physical Sciences Research Council of the UK under grant GR/M42206.

Received 29 March 1999, revised 17 December 1999; published 9 February 2000.

2000 Mathematics Subject Classification 65L05, 65M06

© 2000, Graeme J. Barclay, David F. Griffiths and Desmond J. Higham

a value such as  $\theta = 0.55$  is used as a trade-off between extended stability and second-order accuracy. Exponential fitting [6, 7], the technique whereby  $\theta$  is chosen so that the numerical and exact solutions coincide when  $g(u) = \lambda u$  for given values of  $\lambda$  and  $\Delta t$ , leads to  $\theta \in [1/2, 1]$  for  $\lambda < 0$ . Liniger [7] also shows that the optimality criterion

$$\min_{\theta} \max_{-\infty < z < 0} |e^z - R(z)|$$

leads to the value  $\theta \approx 0.878$ .

Our aim in this work is to investigate the long-term solutions admitted by the theta method. We focus on the *basins of attraction* of fixed points and the *stability for small  $\Delta t$*  of spurious period two solutions; both topics have received little attention in the literature on numerical dynamics. We then study problem classes where the *influence of the period two solutions on the long-term dynamics* can be established rigorously. The related paper [10], which inspired this work, studied the theta method dynamics from a bifurcation viewpoint. For results about the long-term behaviour of general ODE methods, we recommend [2, 4, 5, 9]. The style of analysis in this work is related to that in [1].

The presentation is organised as follows. In Section 2 we give a lower bound on the basin of attraction for fixed points of the theta method that correspond to stable fixed points of the ODE (1). Section 3 concerns period two solutions and their long-term influence. We show that for small  $\Delta t$  these spurious solutions are generically unstable for  $\theta < \frac{1}{2}$ , and stable for  $\theta > \frac{1}{2}$ . We then focus on certain classes of ODE—positive superlinear, negative superlinear, positive sublinear and negative sublinear. In these cases results about the dynamics of equation (1) are readily found and we are able to quantify the corresponding behaviour of the theta method and make clear the negative effect of the period two solutions. In Section 4 the basic period two stability result is extended to a class of semi-discrete partial differential equations. A summary and some conclusions are given in Section 5.

## 2. Basins of attraction for fixed points

It is clear that  $u_j \equiv \beta$  is a solution to equation (2) if and only if  $g(\beta) = 0$ . In other words, the ODE and the method have precisely the same fixed points. In the terminology of [4, 5], the theta method is therefore said to be *regular*.

If  $g(\beta) = 0$  and  $g'(\beta) < 0$ , then the fixed point of equation (1) is linearly stable. By [5, Theorem 3], the corresponding fixed point of the theta method will be stable when  $\Delta t g'(\beta)$  lies in the linear stability interval  $\{z : |R(z)| < 1\}$ , where the stability function  $R$  is defined in equation (3). Straightforward analysis shows that

$$|R(z)| < 1 \Leftrightarrow \begin{cases} \frac{-2}{1-2\theta} < z < 0, & \text{for } 0 \leq \theta < \frac{1}{2}, \\ z < 0, & \text{for } \theta = \frac{1}{2}, \\ z < 0 \text{ or } z > \frac{2}{2\theta-1}, & \text{for } \frac{1}{2} < \theta \leq 1. \end{cases} \quad (4)$$

We see that for  $\theta \geq \frac{1}{2}$ , if  $u(t) \equiv \beta$  is stable for equation (1) then  $u_j \equiv \beta$  is stable for equation (2), independently of  $\Delta t$ . (This is a consequence of the A-stability property; see, for example, [3].) However, for  $\theta > \frac{1}{2}$ , it is possible for the method to stabilize unstable fixed points of the ODE: if  $g'(\beta) > 0$  and  $\Delta t > 2/(g'(\beta)(2\theta - 1))$  then equation (2) has a stable fixed point. For  $\theta < \frac{1}{2}$ , the method will not stabilize unstable fixed points, and will preserve stable fixed points if the timestep is sufficiently small; namely if  $\Delta t < -2/((g'(\beta)(1 - 2\theta))$ .

In terms of capturing the qualitative behaviour of the ODE, it is also of interest to know when the method will exhibit non-oscillatory, monotonic local convergence to a fixed point.

This requires the condition  $|R(z)| < 1$  to be replaced by  $0 < R(z) < 1$ . It is easily shown that

$$0 < R(z) < 1 \Leftrightarrow -1 < (1 - \theta)z < 0.$$

Hence, if  $g'(\beta) < 0$  the method mimics the non-oscillatory local convergence of the ODE when

$$(1 - \theta)\Delta t |g'(\beta)| < 1. \quad (5)$$

We note, however, that linear stability results are concerned with local attractivity. They deal with the existence of non-empty basins of attraction. It is also of interest to have information about the actual basins of attraction of the fixed points. The next theorem shows that if the condition (5) extends to an interval, then non-oscillatory, monotonic convergence is guaranteed throughout the interval.

**Theorem 1.** *Suppose that  $g \in C^1$  with  $g(\beta) = 0$  and  $g'(\beta) < 0$  for some  $\beta \in \mathbb{R}$ . Let  $I \subseteq \mathbb{R}$  be an open, connected interval containing  $\beta$  such that  $g'(u) < 0$  for all  $u \in I$  and let  $g'_{\text{sup}} := \sup_{x \in I} |g'(x)|$ . If*

$$(1 - \theta)g'_{\text{sup}}\Delta t < 1, \quad (6)$$

*then for any  $u_0 \in I$  there exists a solution sequence of equation (2) in which the iterates  $u_j$  lie on the same side of  $\beta$  and approach  $\beta$  monotonically as  $j$  increases.*

*Proof.* Note that  $g$  has a unique root  $\beta$  in  $I$ .

Consider a general iterate  $u_{j-1} \in I$ . If  $u_{j-1} = \beta$  then, trivially,  $u_j \equiv \beta$  and the result follows. Hence, suppose that  $u_{j-1} \neq \beta$ . We define  $h_{j-1} : \mathbb{R} \mapsto \mathbb{R}$  by

$$h_{j-1}(u) := u - \Delta t \theta g(u) - u_{j-1} - \Delta t(1 - \theta)g(u_{j-1}). \quad (7)$$

By construction, a zero of  $h_{j-1}$  provides a solution to equation (2). We have

$$h_{j-1}(u_{j-1}) = -\Delta t g(u_{j-1}) = -\Delta t (g(u_{j-1}) - g(\beta)) = -\Delta t g'(\xi_{j-1}) (u_{j-1} - \beta), \quad (8)$$

where  $\xi_{j-1} \in I$  and we have used the mean-value theorem. Similarly,

$$h_{j-1}(\beta) = \beta - u_{j-1} - \Delta t(1 - \theta)g(u_{j-1}) = (\beta - u_{j-1}) (1 + \Delta t(1 - \theta)g'(\xi_{j-1})). \quad (9)$$

So, equations (8) and (9) give

$$h_{j-1}(u_{j-1})h_{j-1}(\beta) = \Delta t (\beta - u_{j-1})^2 g'(\xi_{j-1}) (1 + \Delta t(1 - \theta)g'(\xi_{j-1})) < 0,$$

where we have used the timestep restriction (6). Hence, there is a zero of  $h_{j-1}$  between  $u_{j-1}$  and  $\beta$ .

The result follows by using compactness and monotonicity.  $\square$

Note that the conditions in Theorem 1 guarantee that for any  $u_0 \in I$ ,  $u(t) \rightarrow \beta$  monotonically as  $t \rightarrow \infty$ . The result establishes a timestep constraint under which the numerical method mimics this behaviour.

When  $\theta = 1$ , equation (6) imposes no restriction on the timestep. Also, in this case the result is equivalent to [1, Theorem 1] (with  $\Delta t$  replaced by  $\Delta x/a$ ).

**Example 1.** We now illustrate Theorem 1 on the logistic ODE, where  $g(u) = u(1 - u)$ . In this case  $\beta = 1$  is a stable fixed point. Since  $g'(u) = 1 - 2u < 0$  for  $u > \frac{1}{2}$ , we must have  $I \subset (\frac{1}{2}, \infty)$  and  $1 \in I$ . To get the largest possible bound on the basin of non-oscillatory, monotonic attraction, we choose  $I$  to match the initial condition as follows.

If  $\frac{1}{2} < u_0 < 1$  then we may take  $I = (u_0 - \epsilon, 1 + \epsilon)$  for any small  $\epsilon$  such that  $u_0 - \epsilon > \frac{1}{2}$ . Then  $g'_{\text{sup}} = 1 + 2\epsilon$  and equation (6) becomes  $(1 - \theta)(1 + 2\epsilon)\Delta t < 1$ . This condition will be satisfied for some small  $\epsilon$  if  $(1 - \theta)\Delta t < 1$ .

For  $u_0 > 1$  we take  $I = (\frac{1}{2}, u_0)$  for which  $g'_{\text{sup}} = 2u_0 - 1$ . The constraint (6) is then  $(1 - \theta)\Delta t(2u_0 - 1) < 1$ .

Overall, we have the following constraint

$$(1 - \theta)\Delta t < \begin{cases} 1 & \text{for } \frac{1}{2} < u_0 < 1, \\ \frac{1}{2u_0 - 1} & \text{for } 1 < u_0. \end{cases} \quad (10)$$

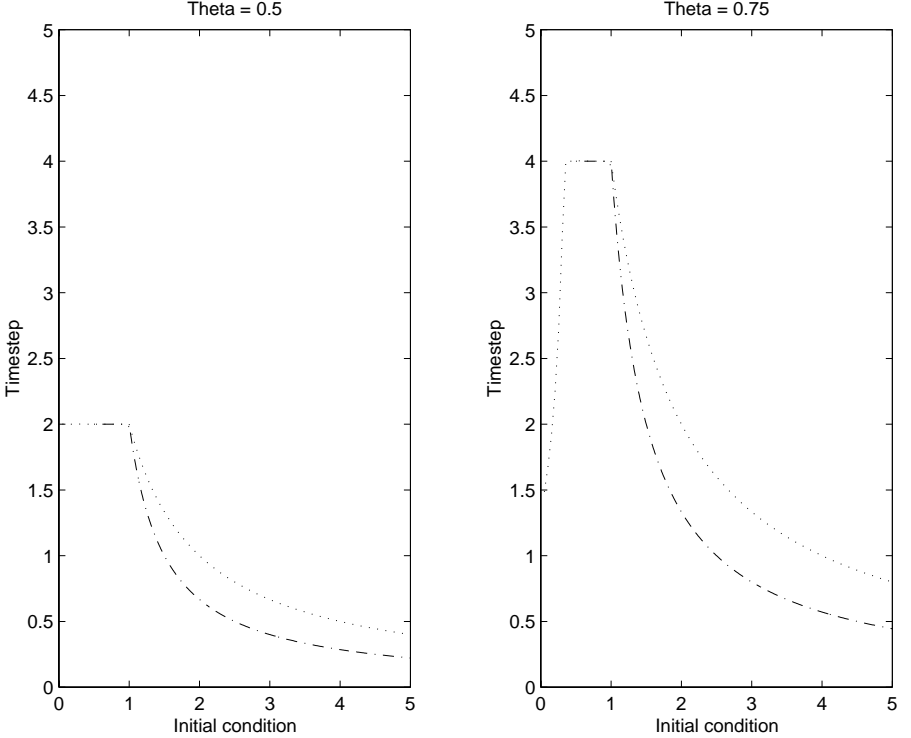


Figure 1: Boundaries for non-oscillatory, monotonic attraction: the dotted line is the computed boundary, while the dash-dotted line is the lower bound

The left- and right-hand pictures in Figure 1 show the constraint defined by inequality (10) in the  $(u_0, \Delta t)$  plane as a dash-dot line in the cases  $\theta = \frac{1}{2}$  and  $\theta = \frac{3}{4}$ , respectively. The dotted line is the corresponding numerically computed constraint. More precisely, the dotted line was computed as follows. For each of a large number of  $u_0$  values, a bisection algorithm was used to compute the largest  $\widehat{\Delta t}(u_0)$  for which non-oscillatory, monotonic convergence was observed for  $0 \leq \Delta t \leq \widehat{\Delta t}(u_0)$ . Non-oscillatory, monotonic convergence was deemed to have occurred if

$$u_j \in \mathbb{R}, \quad (u_{j-1} - \beta)(u_j - \beta) \geq 0, \quad |u_j - \beta| < |u_{j-1} - \beta|, \quad 1 \leq j \leq 10^5$$

and  $|u_{10^5} - \beta| \leq 10^{-3}$ . In solving the quadratic polynomial (2) for  $u_j$ , we took the root closest to  $u_{j-1}$ .

We see from Figure 1 that the constraint (10) does indeed give a lower bound on the region of convergence, and the bound is fairly sharp in this example, especially for  $u_0 \approx \beta$ .

We have already observed from (4) that the theta method with  $\theta > \frac{1}{2}$  may stabilize fixed points that are unstable for the underlying ODE. The following result, which is proved in a different context in [1], applies to the case  $\theta = 1$  and shows how the stabilizing condition  $g'(\beta)\Delta t > 2$  can be generalized to give a lower bound on the basin of attraction.

**Theorem 2.** *Suppose that  $g \in C^1$  with  $g(\beta) = 0$  and  $g'(\beta) > 0$  for some  $\beta \in \mathbb{R}$ . Let  $I \subseteq \mathbb{R}$  be an open, connected interval containing  $\beta$  such that  $g'(u) > 0$  for all  $u \in I$  and let  $g'_{\inf} := \inf_{x \in I} g'(x)$ . Suppose that  $\theta = 1$  and*

$$\Delta t g'_{\inf} > 2. \quad (11)$$

*Then if  $u_0, u_1 \in I$ , there exists a solution sequence of equation (2) in which  $u_j$  approaches  $\beta$  as  $j$  increases, with successive components lying on opposite sides of  $\beta$ .*

*Proof.* The result follows immediately from [1, Theorem 2]. □

### 3. Period two solutions and blow up

Although the theta method never generates spurious fixed points, it is known that spurious period two solutions are admitted. It is demonstrated forcibly in [10] that such solutions, whether stable or unstable, can have a dramatic impact on the long-term dynamics. In this section we prove general results about the nature of period two solutions and, for certain problem classes, quantify precisely their effect on the dynamics.

If  $(v, w)$ , with  $v \neq w$ , is a period two solution of (2), then

$$\begin{aligned} w &= v + \Delta t(1 - \theta)g(v) + \Delta t\theta g(w), \\ v &= w + \Delta t(1 - \theta)g(w) + \Delta t\theta g(v). \end{aligned}$$

These conditions are equivalent to

$$v + \Delta t(1 - 2\theta)g(v) = w, \quad (12)$$

$$g(v) + g(w) = 0. \quad (13)$$

It follows immediately from equation (12) (and is shown in [10]) that period two solutions cannot exist for  $\theta = \frac{1}{2}$ .

Writing the theta method as  $u_j = S(u_{j-1})$ , we have

$$S(u) = u + (1 - \theta)\Delta t g(u) + \theta \Delta t g(S(u)).$$

Hence, if  $1 - \theta \Delta t g'(S(u)) \neq 0$ ,

$$S'(u) = \frac{1 + (1 - \theta)\Delta t g'(u)}{1 - \theta \Delta t g'(S(u))}. \quad (14)$$

A period two solution  $(v, w)$  is linearly stable if  $|S'(v)S'(w)| < 1$  and linearly unstable if  $|S'(v)S'(w)| > 1$ . Since, by definition,  $v = S(w)$  and  $w = S(v)$ , it follows that these conditions may be written as

$$|R(z_v)R(z_w)| < 1 \quad \text{and} \quad |R(z_v)R(z_w)| > 1, \quad (15)$$

respectively, where  $z_v = \Delta t g'(v)$  and  $z_w = \Delta t g'(w)$  and the stability function  $R$  is defined in equation (3).

**Example 2.** For  $\theta < \frac{1}{2}$ , an example of a period two solution is given by

$$g(u) = -2u|u|, \quad v = \frac{1}{(2\theta - 1)\Delta t}, \quad w = -v. \quad (16)$$

Similarly, for  $\theta > \frac{1}{2}$

$$g(u) = 2u|u|, \quad v = \frac{1}{(2\theta - 1)\Delta t}, \quad w = -v, \quad (17)$$

defines a period two solution. In both cases we find that

$$g'(v) = g'(w) = \frac{4}{(2\theta - 1)\Delta t}, \quad \text{and} \quad |R(z_v)R(z_w)| = \left| \frac{3 - 2\theta}{1 + 2\theta} \right|^2. \quad (18)$$

Hence, these period two solutions are linearly stable if and only if  $\theta > \frac{1}{2}$ .

This example has a number of features of interest: the period two solutions blow up monotonically and in opposite directions as  $\Delta t \rightarrow 0$ , the derivative values  $g'(v)$  and  $g'(w)$  blow up in the same direction as  $\Delta t \rightarrow 0$ , and there is a change in stability as  $\theta$  crosses  $\frac{1}{2}$ . In the analysis below we show that, with mild assumptions on  $g$ , these features can be shown to be generic.

In Lemma 3 below, we show that genuine period two solutions must exhibit a precise form of blow up as  $\Delta t \rightarrow 0$ . (We note that the fact that  $|v|, |w| \rightarrow \infty$  was proved by a different approach in [10] and also follows from the general theory of Humphries [4].)

**Lemma 3.** *Consider the theta method applied to the scalar ODE (1), where  $g$  is continuous. Suppose that for sufficiently small  $\Delta t$ , there is a period two solution  $v = v(\Delta t)$  and  $w = w(\Delta t)$  with  $v$  and  $w$  depending continuously upon  $\Delta t$  and with  $g(v)$  and  $g(w)$  bounded away from zero for small  $\Delta t$ . Then as  $\Delta t \rightarrow 0$*

$$|v| \rightarrow \infty, \quad |w| \rightarrow \infty, \quad |g(v)| \rightarrow \infty, \quad |g(w)| \rightarrow \infty.$$

Further, for small  $\Delta t$ ,

$$vw < 0, \quad (1 - 2\theta)vg(v) < 0, \quad (1 - 2\theta)wg(w) < 0,$$

and  $v$  and  $w$  blow up monotonically.

*Proof.* Recall that  $u$  and  $v$  must satisfy equations (12) and (13).

If  $v$  is bounded as  $\Delta t \rightarrow 0$  then  $\Delta t g(v) \rightarrow 0$ , so that  $w \rightarrow v$  in equation (12). In this case, from equation (13),  $g(v) \rightarrow 0$ . This contradicts the assumption that  $g(v)$  is bounded away from zero for small  $\Delta t$ , and hence we must have  $|v| \rightarrow \infty$  as  $\Delta t \rightarrow 0$ . Similarly,  $|w| \rightarrow \infty$ . Recalling that  $g(v)$  and  $g(w)$  are bounded away from zero, it follows from equation (13) that  $vw < 0$  for small  $\Delta t$ . Subsequently, using equation (12),

$$|g(v)| = \frac{|v| + |w|}{\Delta t |1 - 2\theta|} \rightarrow \infty, \quad \text{as } \Delta t \rightarrow 0.$$

Similarly,  $|g(w)| \rightarrow \infty$  as  $\Delta t \rightarrow 0$ .

From equation (12) we have, for small  $\Delta t$ ,

$$v^2 + \Delta t(1 - 2\theta)vg(v) = vw < 0$$

and so  $(1 - 2\theta)vg(v) < 0$ . By symmetry,  $(1 - 2\theta)wg(w) < 0$ .

Before proving monotonic blow up, we first show that

$$\liminf |v| = \infty \quad \text{and} \quad \liminf |w| = \infty, \quad (19)$$

where  $\liminf(\cdot)$  means  $\lim_{\Delta t^* \rightarrow 0}(\inf_{0 < \Delta t < \Delta t^*}(\cdot))$ . For contradiction, suppose without loss of generality that  $\liminf |v| < \infty$ . Then, in equation (12),

$$\liminf |v - w| \leq |1 - 2\theta| \liminf \Delta t |g(v)| = 0.$$

By continuity,  $\liminf |g(v) - g(w)| = 0$ , and hence, using equation (13),  $\liminf |g(v)| = 0$ . This gives the required contradiction.

Finally, we show by contradiction that  $v$  and  $w$  blow up monotonically for small  $\Delta t$ . We assume without loss of generality that  $v \rightarrow \infty$  and  $w \rightarrow -\infty$  as  $\Delta t \rightarrow 0$ . Suppose that  $v$  loses monotonicity at some small timestep  $\Delta t^*$ . Then, if, at any timestep  $\widehat{\Delta t} < \Delta t^*$ ,  $v(\widehat{\Delta t}) = v(\Delta t^*)$ , it follows from equation (12) that  $w(\widehat{\Delta t}) > w(\Delta t^*)$ . Similarly, if, at any timestep  $\widehat{\Delta t} < \Delta t^*$ ,  $w(\widehat{\Delta t}) = w(\Delta t^*)$ , it follows from equation (12) that  $v(\widehat{\Delta t}) < v(\Delta t^*)$ . This contradicts equations (19).  $\square$

We now use this result to make general conclusions about the stability of period two solutions. Note that the theorem below applies to the example (16)–(17) and to [10, Examples 2.1 and 2.2].

**Theorem 4.** *Under the conditions of Lemma 3, if  $g \in C^1$  then*

1. *for  $\theta < \frac{1}{2}$  the period two solution is unstable for small  $\Delta t$ , and*
2. *for  $\theta > \frac{1}{2}$  the period two solution is stable for small  $\Delta t$ .*

*Proof.* Consider first the case  $\theta < \frac{1}{2}$ . Appealing to Lemma 3, and assuming without loss of generality that  $v \rightarrow +\infty$  rather than  $-\infty$ , we have

$$v \rightarrow \infty, \quad g(v) \rightarrow -\infty, \quad w \rightarrow -\infty, \quad g(w) \rightarrow \infty.$$

Now consider a fixed, small  $\Delta t > 0$  with corresponding  $v$  and  $w$ . Perturb  $\Delta t$  to  $\Delta t - \delta t > 0$ , where  $\delta t > 0$  is small. This gives us  $v + \delta v$  and  $w + \delta w$  and, from the monotonicity result in Lemma 3, we have  $\delta v > 0$  and  $\delta w < 0$ .

Note from equation (12) that

$$\frac{g(v)}{v - w} = \frac{-1}{\Delta t(1 - 2\theta)} \quad (20)$$

always holds. Hence reducing  $\Delta t$  towards zero causes the left-hand side of equation (20) to become more negative; that is,

$$\frac{g(v + \delta v)}{v + \delta v - (w + \delta w)} < \frac{g(v)}{v - w}.$$

Hence, for sufficiently small perturbations,

$$(g(v) + \delta v g'(v))(v - w) < g(v)(v - w + \delta v - \delta w),$$

and so

$$\delta v g'(v)(v - w) < g(v)(\delta v - \delta w) = (v - w) \frac{-1}{\Delta t(1 - 2\theta)} (\delta v - \delta w),$$

where we have once more made use of equation (20). This simplifies to

$$g'(v) < \frac{-1}{\Delta t(1-2\theta)} \left(1 - \frac{\delta w}{\delta v}\right). \quad (21)$$

By symmetry, we find that

$$g'(w) < \frac{-1}{\Delta t(1-2\theta)} \left(1 - \frac{\delta v}{\delta w}\right). \quad (22)$$

Recall that the conditions in inequalities (15) determine stability, and note that

$$R(z_v)R(z_w) - 1 = \frac{\Delta t g'(v) + \Delta t g'(w) + (1-2\theta)\Delta t g'(v)\Delta t g'(w)}{[1 - \theta \Delta t g'(v)][1 - \theta \Delta t g'(w)]}.$$

It is straightforward to show that both the numerator and denominator are strictly positive if inequalities (21) and (22) hold. Hence,  $|R(z_v)R(z_w)| > 1$ , as required.

For  $\theta > \frac{1}{2}$  we may assume without loss of generality that

$$v \rightarrow \infty, \quad g(v) \rightarrow \infty, \quad w \rightarrow -\infty, \quad g(w) \rightarrow -\infty,$$

and a similar analysis to that above can be performed, resulting in the inequalities

$$g'(v) > \frac{1}{\Delta t(2\theta-1)} \left(1 - \frac{\delta w}{\delta v}\right) \quad (23)$$

$$g'(w) > \frac{1}{\Delta t(2\theta-1)} \left(1 - \frac{\delta v}{\delta w}\right). \quad (24)$$

It can then be shown that  $|R(z_v)R(z_w)| < 1$ , giving the required result.  $\square$

We now study four problem classes for which the existence of period two solutions and their effect on the long-term dynamics can be pinned down precisely.

### Definitions

- Function  $g$  is *positive superlinear* if  $g \in C^1$  and  $g'(u) \rightarrow +\infty$  as  $|u| \rightarrow \infty$ .
- Function  $g$  is *negative superlinear* if  $g \in C^1$  and  $g'(u) \rightarrow -\infty$  as  $|u| \rightarrow \infty$ .
- Function  $g$  is *positive sublinear* if  $g \in C^1$  and there exists a constant  $D$  such that  $0 < g'(u) < D$  for all  $u$ .
- Function  $g$  is *negative sublinear* if  $g \in C^1$ , there exists a constant  $D$  such that  $-D < g'(u) < 0$  for all  $u$  and  $g(u^*) = 0$  for some  $u^* \in (-\infty, \infty)$ . In this case  $u^*$  is unique, and we assume for convenience that  $u^* = 0$ .

The following results are easily established.

### Results

Suppose that equation (1) has a unique solution for all  $u_0 \in \mathbb{R}$  and  $t > 0$ .

- If  $g$  is positive superlinear then for  $|u_0|$  sufficiently large in equation (1),  $|u(t)| \rightarrow \infty$  monotonically as  $t \rightarrow \infty$ .
- If  $g$  is negative superlinear then equation (1) is dissipative in the sense that there exists a constant  $K$  such that  $|u(t)|$  decreases monotonically with  $t$  whenever  $|u(t)| > K$ .
- If  $g$  is positive sublinear then whenever  $g(u_0) \neq 0$  in equation (1),  $u(t)$  is monotonic and  $|u(t)| \rightarrow \infty$  as  $t \rightarrow \infty$ .
- If  $g$  is negative sublinear then all solutions to equation (1) satisfy  $u(t) \rightarrow 0$  as  $t \rightarrow \infty$ .



Theorem 5 below shows that in the positive superlinear case, period two solutions exist if and only if  $\theta > \frac{1}{2}$ , and the monotonic asymptotic blow up property of the problem (1) is not captured when  $\theta \neq 0$ .

**Theorem 5.** *Consider the theta method applied to the scalar ODE (1), where  $g$  is positive superlinear:*

1. *If and only if  $\theta > \frac{1}{2}$  do there exist for small  $\Delta t$  period two solutions  $v = v(\Delta t)$  and  $w = w(\Delta t)$  with  $v$  and  $w$  depending continuously upon  $\Delta t$  and with  $g(v)$  and  $g(w)$  bounded away from zero for small  $\Delta t$ . (Note that Lemma 3 and Theorem 4 apply to these solutions.)*
2. *For  $\theta \neq 0$  there does not exist a numerical solution such that  $|u_j| \rightarrow \infty$  as  $j \rightarrow \infty$  and  $\{u_j\}_{j=0}^\infty$  is monotonic.*

*Proof.* We begin by proving the ‘if’ implication of part 1. Suppose that  $\theta > \frac{1}{2}$ . We note that since  $g$  is superlinear,  $g(u)$  is monotonic for large  $|u|$ , say  $|u| \geq L$ . We may redefine  $g(u)$  for  $|u| < L$  without affecting the validity of the proof, and hence we suppose that  $g(u)$  is monotonic for all  $u$ .

It follows from equations (12) and (13) that  $v, w$  with  $v \neq w$  is a period two solution if

$$\frac{g(v) - g(w)}{v - w} = \frac{1}{\Delta t(\frac{\theta}{2} - 1)}, \quad (25)$$

$$g(v) = -g(w). \quad (26)$$

Now consider the straight line through the origin of (positive) slope  $1/(\Delta t(\frac{\theta}{2} - 1))$ . Since  $g$  is superlinear, for small  $\Delta t$  there must be points  $v > 0$  and  $w < 0$  at which this line intersects  $g$ , so that equation (25) is satisfied. Now, since  $g$  is monotonic, by adding a constant to the straight-line function we may alter the intersection points until equation (26) is satisfied. This establishes the existence of a period two solution for all small  $\Delta t$ . The solution is clearly continuous in  $\Delta t$ .

To prove the ‘only if’ implication of part 1, suppose that  $\theta < \frac{1}{2}$ . Note that  $\text{sgn}(g(u)) = \text{sgn}(u)$  for large  $|u|$ , say  $|u| \geq M$ . We may redefine  $g(u)$  for  $|u| < M$  without affecting the validity of the proof, and hence we suppose that  $\text{sgn}(g(u)) = \text{sgn}(u)$  for all  $u \neq 0$ . By Lemma 3, if  $(v, w)$  is a period two solution then  $vw < 0$ , so that  $(g(v) - g(w))/(v - w) > 0$ , which contradicts equation (25).

To prove part 2 we let  $p : \mathbb{R} \rightarrow \mathbb{R}$  be defined by  $p(u) = u - \theta \Delta t g(u)$ . The theta method (2) may then be written

$$p(u_j) - p(u_{j-1}) = \Delta t g(u_{j-1}).$$

From the mean-value theorem,

$$p'(z_j)(u_j - u_{j-1}) = \Delta t g(u_{j-1}),$$

where  $z_j$  lies between  $u_{j-1}$  and  $u_j$ . This means that

$$(1 - \theta \Delta t g'(z_j))(u_j - u_{j-1}) = \Delta t g(u_{j-1}). \quad (27)$$

If  $|u_{j-1}|$  and  $|u_j|$  are sufficiently large and of the same sign, then  $1 - \theta \Delta t g'(z_j) < 0$  in equation (27), and the result follows.  $\square$

Part 1 of Theorem 5 shows that stable period two solutions are generic for  $\theta > \frac{1}{2}$  on positive superlinear blow up problems. The proof of part 2 shows that the solution sequence

will increase monotonically until it reaches a point where  $1 - \theta \Delta t g'(u) < 0$ , after which it may be expected to approach the stable period two solution. We illustrate this behaviour in the case  $g(u) = u \ln(1 + |u|)$  with  $u_0 = 1$  and  $\Delta t = 0.1$ . The upper and lower pictures in Figure 2 show  $\{u_j\}_{j=0}^{300}$  for  $\theta = \frac{3}{4}$  and  $\theta = 1$ , respectively. For clarity, we plot the piecewise linear interpolant through the data. The dashed lines show period two solutions that were found by solving equations (12) and (13). In both cases, the numerical solution increases monotonically until the condition  $1 - \theta \Delta t g'(u_j) > 0$  is first violated. From this point onwards the solution approaches the stable period two level.

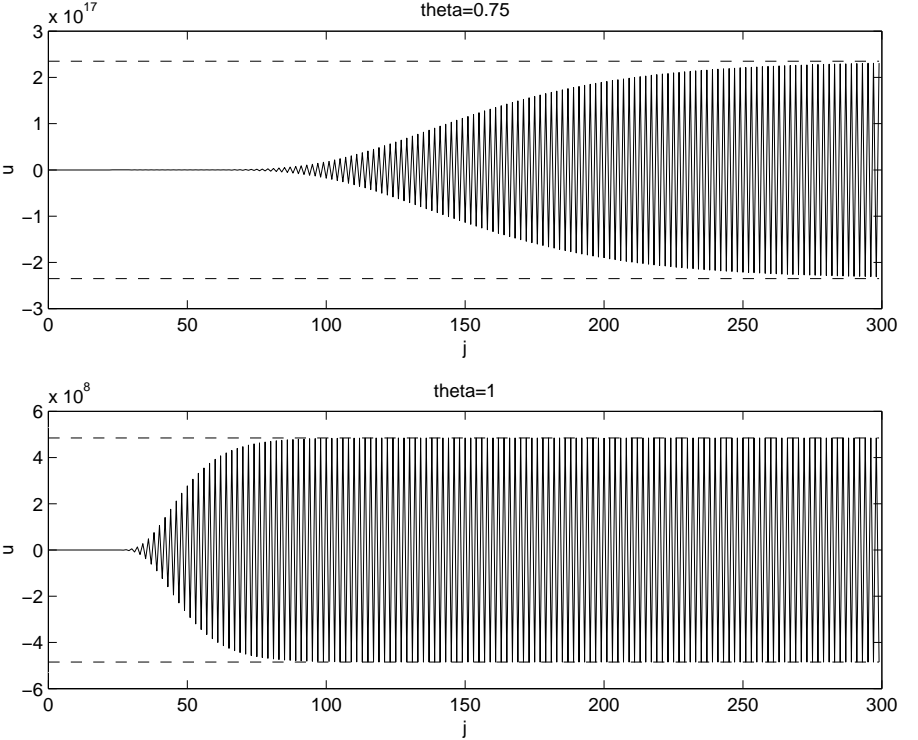


Figure 2: Theta method solutions with  $\theta = \frac{3}{4}$  (upper) and  $\theta = 1$  (lower) for  $g(u) = u \ln(1 + |u|)$  with  $u_0 = 1$  and  $\Delta t = 0.1$ .

We remark that a related area—finite time blow up—has been studied by Sanz-Serna and Verwer [8]; for  $\theta = 0$ , they took  $g(u) = u^m$  (for  $m > 1$ ) and  $u_0 = 1$ . In this case, solutions of problem (1) exist only for  $0 < t < 1/(m - 1)$ , and it is shown that Euler’s method mimics the correct behaviour as  $\Delta t \rightarrow 0$ .

Theorem 6 below concerns the negative superlinear case. It shows that period two solutions exist if and only if  $\theta < \frac{1}{2}$  and in this parameter range the dissipativity property of the problem (1) is not captured.

**Theorem 6.** Consider the theta method applied to the scalar ODE (1), where  $g$  is negative superlinear.

1. If and only if  $\theta < \frac{1}{2}$  do there exist for small  $\Delta t$  period two solutions  $v = v(\Delta t)$  and  $w = w(\Delta t)$  with  $v$  and  $w$  depending continuously upon  $\Delta t$  and with  $g(v)$  and  $g(w)$  bounded away from zero for small  $\Delta t$ . (Note that Lemma 3 and Theorem 4 apply to these solutions.)
2. If  $\theta < \frac{1}{2}$  then the theta method is not dissipative in the sense that there exists a  $K = K(\Delta t, \theta)$  such that for every  $\widehat{K} > K$  there is a pair  $u_0, u_1$  with  $|u_1| > |u_0| = \widehat{K}$  that satisfies equation (2).

*Proof.* Part 1 may be proved in a similar manner to part 1 of Theorem 5.

To prove part 2, let  $\theta = \frac{1}{2} - \epsilon$ , where  $\epsilon > 0$ . For a given  $u_0$ , let  $h_0(u)$  be defined by equation (7), so that

$$h_0(u) := u - \frac{\Delta t}{2} (g(u) + g(u_0)) + \epsilon \Delta t (g(u) - g(u_0)) - u_0.$$

Note that  $h_0(u) \rightarrow \infty$  as  $u \rightarrow \infty$  and  $h_0(u) \rightarrow -\infty$  as  $u \rightarrow -\infty$ .

Since  $g$  is negative superlinear, there is a  $K = K(\Delta t, \theta)$  such that  $g$  is monotonic for  $|u| > K$  and

$$|u| > K \Rightarrow g'(u) < 0, \quad ug(u) < 0, \quad |g(u)| > \frac{|u|}{\epsilon \Delta t}. \quad (28)$$

Now, consider any  $\widehat{u} > K$ . If  $|g(\widehat{u})| \geq |g(-\widehat{u})|$  then it follows from inequalities (28) that, with  $u_0 = \widehat{u}$ ,

$$h_0(-\widehat{u}) = -2\widehat{u} - \frac{\Delta t}{2} (g(-\widehat{u}) + g(\widehat{u})) + \epsilon \Delta t (g(-\widehat{u}) - g(\widehat{u})) > 0.$$

Hence, there is a zero of  $h_0$  in the interval  $(-\infty, -\widehat{u})$ . On the other hand, if  $|g(\widehat{u})| < |g(-\widehat{u})|$  then it follows from inequalities (28) that, with  $u_0 = -\widehat{u}$ ,

$$h_0(\widehat{u}) = 2\widehat{u} - \frac{\Delta t}{2} (g(\widehat{u}) + g(-\widehat{u})) + \epsilon \Delta t (g(\widehat{u}) - g(-\widehat{u})) < 0.$$

Hence, there is a zero of  $h_0$  in the interval  $(\widehat{u}, \infty)$ .

Overall, we have shown that for every  $\widehat{K} > K$  there exists a pair  $u_0, u_1$  with  $|u_1| > |u_0| = \widehat{K}$  that satisfies equation (2).  $\square$

By Theorem 4, the period two solution identified in part 1 of Theorem 6 must be unstable. Hence, it is reasonable to regard the unstable spurious solution as the cause of the non-dissipativity established in part 2. To illustrate this idea, we consider the case where  $g(u) = -u(u+1)(u-1)$ . In this case a period two solution for  $\theta < \frac{1}{2}$  can be found analytically—the positive branch  $\sqrt{1 - 2/(\Delta t(1 - 2\theta))}$  is plotted with the ‘o’ symbol in Figure 3 for  $\theta = \frac{1}{4}$ . The dark and light regions in Figure 3 show the timesteps and initial conditions in the range  $0.2 \leq \Delta t \leq 1$  and  $0 \leq u_0 \leq 5$ , for which the theta method produced dissipative and non-dissipative solutions, respectively. In these computations, a solution was regarded as non-dissipative if  $\max_{0 \leq j \leq 50} |u_j| > 5000$ . We see that the unstable period two branch clearly delimits the correct and incorrect asymptotic behaviour.

Lemma 7 below shows that the result in part 2 of Theorem 6 does not extend to  $\theta \geq \frac{1}{2}$ , and in this sense the cut-off for a guaranteed lack of dissipativity coincides with the cut-off for the existence of period two solutions.

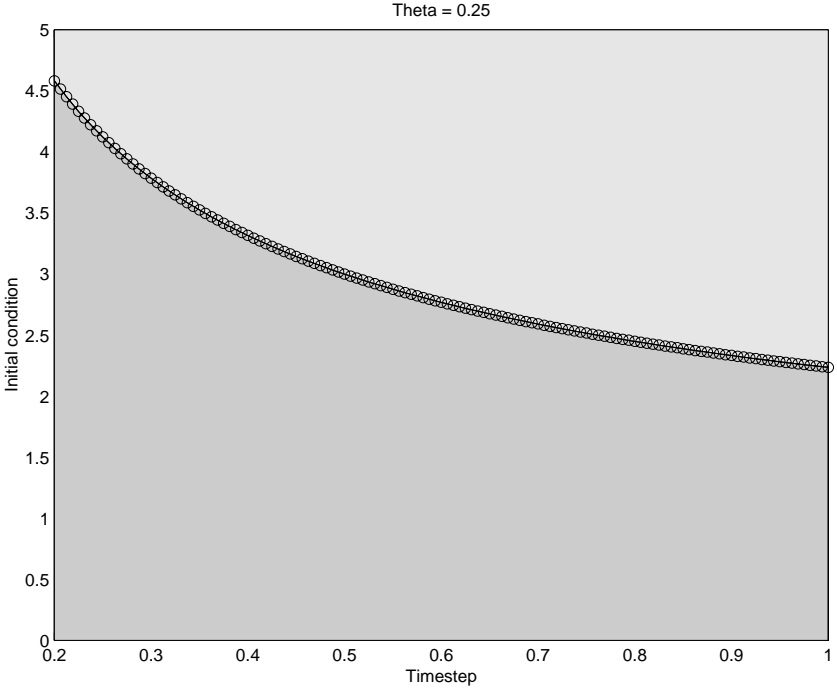


Figure 3: Theta method with  $\theta = \frac{1}{4}$  on  $g(u) = -u(u + 1)(u - 1)$ . The dark region leads to dissipative solutions, the light region leads to non-dissipative solutions, and ‘o’ marks the unstable period two branch.

**Lemma 7.** Suppose that  $g \in C^1$  is odd and  $g'(u) < 0$  for all  $u$ . Then, for  $\theta \geq \frac{1}{2}$ , any solution sequence  $\{u_j\}_{j=0}^\infty$  produced by the theta method has  $|u_j| \rightarrow 0$  monotonically as  $j \rightarrow \infty$ .

*Proof.* First, we note that  $u_{j-1} = 0 \Rightarrow u_j = 0$ .

Now, suppose that  $u_{j-1} > 0$ . If  $u_j > 0$  then  $g(u_j) < 0$  and  $g(u_{j-1}) < 0$ , and hence

$$u_j = u_{j-1} + \Delta t (\theta g(u_j) + (1 - \theta)g(u_{j-1})) < u_{j-1}.$$

On the other hand, if  $u_j < 0$  then suppose that  $|u_j| > |u_{j-1}|$ . We then have  $g(u_j) > -g(u_{j-1})$ , and hence  $\theta g(u_j) > -(1 - \theta)g(u_{j-1})$ . This gives

$$0 > u_j = u_{j-1} + \Delta t (\theta g(u_j) + (1 - \theta)g(u_{j-1})) > u_{j-1} > 0,$$

which is a contradiction.

We have thus shown that

$$u_{j-1} > 0 \Rightarrow |u_j| < |u_{j-1}|.$$

Similarly, we can show that

$$u_{j-1} < 0 \Rightarrow |u_j| < |u_{j-1}|.$$

Hence  $|u_j| < |u_{j-1}|$  whenever  $u_{j-1} \neq 0$ .

Since the sequence  $\{|u_j|\}_{j=0}^{\infty}$  is bounded, it must have a convergent subsequence, with limit  $u^*$ . It follows from equation (2) that  $g(u^*) = 0$ , and hence that  $u^* = 0$ . The monotonicity result then shows that the full sequence  $\{|u_j|\}_{j=0}^{\infty}$  must have the same limit.  $\square$

Theorem 8 below concerns the positive sublinear case. It shows that for small  $\Delta t$ , period two solutions do not exist and monotonic asymptotic blow up is guaranteed.

**Theorem 8.** *Consider the theta method applied to the scalar ODE (1), where  $g$  is positive sublinear.*

1. *Period two solutions do not exist for small  $\Delta t$ .*
2. *If  $g(u_0) \neq 0$  then for sufficiently small  $\Delta t$ , any numerical solution sequence  $\{u_j\}_{j=0}^{\infty}$  is monotonic and satisfies  $|u_j| \rightarrow \infty$  as  $j \rightarrow \infty$ .*

*Proof.* Note that if  $(v, w)$  is a period two solution then  $\theta \neq \frac{1}{2}$  and, applying the mean-value theorem to equation (25),

$$g'(z) = \frac{1}{\Delta t(\frac{\theta}{2} - 1)}, \quad (29)$$

where  $z$  lies between  $w$  and  $v$ . Since  $g$  is sublinear, this cannot hold when  $\Delta t$  is sufficiently small. This proves part 1.

To prove part 2 we follow the proof in Theorem 5, part 2, and note that  $1 - \theta \Delta t g'(z_j) > 0$  in equation (27) for small  $\Delta t$ .  $\square$

Theorem 9 below concerns the negative sublinear case. It shows that for small  $\Delta t$ , period two solutions do not exist, and convergence to a steady state is guaranteed.

**Theorem 9.** *Consider the theta method applied to the scalar ODE (1), where  $g$  is negative sublinear.*

1. *Period two solutions do not exist for small  $\Delta t$ .*
2. *For sufficiently small  $\Delta t$ , any numerical solution sequence  $\{u_j\}_{j=0}^{\infty}$  is monotonic and satisfies  $u_j \rightarrow 0$  as  $j \rightarrow \infty$ .*

*Proof.* Part 1 follows from equation (29).

To prove part 2 we consider the nontrivial case  $u_{j-1} \neq 0$ . Since  $g$  is sublinear and  $g(0) = 0$ ,

$$|g(u)| \leq D|u|, \quad \text{for all } u. \quad (30)$$

Let

$$\Delta t \leq \frac{1}{2D}. \quad (31)$$

From equations (27), (30) and (31) we have

$$|u_j - u_{j-1}| = \left| \frac{\Delta t g(u_{j-1})}{1 - \theta \Delta t g'(z_j)} \right| \leq \Delta t |g(u_{j-1})| \leq \frac{|u_{j-1}|}{2}. \quad (32)$$

We deduce that  $u_j$  and  $u_{j-1}$  always have the same sign. It follows from equation (27) that  $\text{sgn}(u_j - u_{j-1}) = \text{sgn}(g(u_{j-1})) = -\text{sgn}(u_{j-1})$ . We conclude that  $|u_j| < |u_{j-1}|$ , which completes the proof.  $\square$

4. Partial differential equations

We now generalize Theorem 4 to cover a class of semi-discretized partial differential equations. More precisely, we consider ODE systems of the form

$$\mathbf{U}' = -\frac{a}{\Delta x^m} \mathbf{A} \mathbf{U} + \mathbf{g}(\mathbf{U}) =: G(\mathbf{U}), \quad \mathbf{U}(0) = \mathbf{U}_0 \in \mathbb{R}^N, \quad (33)$$

where  $a, \Delta x > 0$ ,  $m \in \mathbb{Z}^+$ ,  $A \in \mathbb{R}^{N \times N}$  and  $\mathbf{g}(\mathbf{U})_i \equiv g(U_i)$  with  $g : \mathbb{R} \mapsto \mathbb{R}$ . Such systems arise, for example, when a method-of-lines approach is used to solve a periodic, initial-value problem that combines reaction with convection or diffusion in one or more space variables. Here, the spatial derivatives are discretized using finite differences or finite elements, with  $\Delta x$  representing the spatial mesh size.

The theta method applied to system (33) takes the form

$$\mathbf{U}_j = \mathbf{U}_{j-1} + \Delta t(1 - \theta)G(\mathbf{U}_{j-1}) + \Delta t\theta G(\mathbf{U}_j).$$

We will suppose that the matrix  $A$  in system (33) satisfies  $\mathbf{A}\mathbf{e} = \mathbf{0}$ , where  $\mathbf{e} \in \mathbb{R}^N$  is the vector with all components equal to 1. In this case, fixed points or periodic solutions of the theta method on the scalar problem (1) correspond to spatially uniform fixed points or periodic solutions of the theta method on the system (33).

**Theorem 10.** *Consider the theta method applied to the system (33), where  $\mathbf{A}\mathbf{e} = \mathbf{0}$  and  $g \in C^1$ . Consider a spatially uniform period two solution  $\{\mathbf{v}\mathbf{e}, \mathbf{w}\mathbf{e}\}$ , where  $v = v(\Delta t)$  and  $w = w(\Delta t)$  with  $v$  and  $w$  depending continuously upon  $\Delta t$  and with  $g(v)$  and  $g(w)$  bounded away from zero for small  $\Delta t$ . Let  $c := a\Delta t / \Delta x^m$ .*

1. For  $\theta < \frac{1}{2}$ , independently of  $c$ , this period two solution is linearly unstable for small  $\Delta t$ .
2. For  $\theta > \frac{1}{2}$ , this period two solution is linearly stable for small  $c$ .

*Proof.* Writing system (33) as  $\mathbf{U}' = G(\mathbf{U})$ , the Jacobian of  $G$  at a point  $u\mathbf{e}$  has the form

$$G'(u\mathbf{e}) = -\frac{a}{\Delta x^m} A + g'(u)I. \quad (34)$$

Writing the theta method as  $\mathbf{U}_j = S(\mathbf{U}_{j-1})$ , the Jacobian of  $S$  at a point  $u\mathbf{e}$  has the form

$$S'(u\mathbf{e}) = [I - \theta\Delta t G'(S(u\mathbf{e}))]^{-1} [I + (1 - \theta)\Delta t G'(u\mathbf{e})],$$

and hence

$$\begin{aligned} S'(v\mathbf{e})S'(w\mathbf{e}) &= [I - \theta\Delta t G'(w\mathbf{e})]^{-1} [I + (1 - \theta)\Delta t G'(v\mathbf{e})] \\ &\quad \times [I - \theta\Delta t G'(v\mathbf{e})]^{-1} [I + (1 - \theta)\Delta t G'(w\mathbf{e})]. \end{aligned} \quad (35)$$

If  $A$  has an eigenvalue  $\lambda$  then, from equations (34) and (35),  $S'(v\mathbf{e})S'(w\mathbf{e})$  has an eigenvalue

$$\frac{(1 + (1 - \theta)\Delta t(-\frac{a}{\Delta x^m}\lambda + g'(v))) (1 + (1 - \theta)\Delta t(-\frac{a}{\Delta x^m}\lambda + g'(w)))}{(1 - \theta\Delta t(-\frac{a}{\Delta x^m}\lambda + g'(v))) (1 - \theta\Delta t(-\frac{a}{\Delta x^m}\lambda + g'(w)))}, \quad (36)$$

which may be written  $R(\widehat{z}_v)R(\widehat{z}_w)$ , where

$$\widehat{z}_v = \Delta t(-\frac{a}{\Delta x^m}\lambda + g'(v)) \quad \text{and} \quad \widehat{z}_w = \Delta t(-\frac{a}{\Delta x^m}\lambda + g'(w)). \quad (37)$$

For  $\theta < \frac{1}{2}$ , inserting the eigenvalue  $\lambda = 0$ , and following the proof of Theorem 4 establishes instability for small  $\Delta t$ .

For  $\theta > \frac{1}{2}$ , we know from Theorem 4 that  $|R(\Delta t g'(v))R(\Delta t g'(w))| < 1$  for small  $\Delta t$ . Since  $\widehat{z}_v = \Delta t g'(v) + O(c)$  and  $\widehat{z}_w = \Delta t g'(w) + O(c)$ , the required stability result follows.  $\square$

It is of interest to contrast Theorem 10 with [1, Theorem 6]. In [1] a class of explicit Runge–Kutta methods is studied, and spurious fixed points (of period one) are considered. It is shown that although stable spurious solutions can exist for small  $\Delta t$  on a scalar ODE, these must necessarily be unstable as spatially uniform spurious fixed points of a method-of-lines system (33). The theorem above, on the other hand, concerns a different class of time-stepping methods, and involves period two solutions. The result shows that for  $\theta > \frac{1}{2}$  stable spurious behaviour is generic for small  $c$ .

The following example, to which Theorem 10 applies, is sufficiently simple that we can compute precise stability constraints.

**Example 3.** Consider the system (33) with  $g(u)$  from equations (16)–(17),  $m = 1$  and

$$A = \begin{bmatrix} 1 & & -1 \\ -1 & 1 & \\ & \ddots & \ddots \\ & & -1 & 1 \end{bmatrix}. \quad (38)$$

This system arises when first-order upwind differences are used on the hyperbolic problem  $u_t + au_x = g(u)$ , with  $u(x, 0)$  given for  $0 < x < 1$  and periodic boundary conditions. The eigenvalues of  $A$  are

$$\lambda_k = 1 - \exp\left(\frac{2\pi ik}{N}\right), \quad 1 \leq k \leq N. \quad (39)$$

We note that for  $z \in \mathbb{C}$ , writing  $z = x + iy$ ,

$$|R(x + iy)| < 1 \Leftrightarrow \begin{cases} \left(x + \frac{1}{1-2\theta}\right)^2 + y^2 < \frac{1}{(1-2\theta)^2}, & \text{for } 0 \leq \theta < \frac{1}{2}, \\ \left(x + \frac{1}{1-2\theta}\right)^2 + y^2 > \frac{1}{(1-2\theta)^2}, & \text{for } \frac{1}{2} < \theta \leq 1. \end{cases} \quad (40)$$

When  $\theta < \frac{1}{2}$  we consider  $k = N$  in equations (39). In this case, from equation (18), we have  $\widehat{z}_v = \widehat{z}_w = \frac{4}{2\theta-1}$  which does not lie in the region given by implication (40). Hence, the solution is always unstable.

To study the case  $\theta > \frac{1}{2}$ , we note from equation (18) that  $\widehat{z}_v$  and  $\widehat{z}_w$  in equations (37) lie on the circle

$$\left(x + c - \frac{4}{2\theta-1}\right)^2 + y^2 = c^2, \quad (41)$$

where  $c := a\Delta t/\Delta x$  is the Courant number. Comparing this with (40), we find that, as  $c$  increases, stability can be guaranteed until the circle (41) intersects the circle

$$\left(x + \frac{1}{1-2\theta}\right)^2 + y^2 = \frac{1}{(1-2\theta)^2}.$$

It follows that the period two solution is linearly stable if  $c < \frac{1}{2\theta-1}$ .

## 5. Summary and conclusions

The popularity of the theta method is due in large part to its simplicity—making it (a) easy to program and (b) efficient on large problems. In this work we aim to show that

the simple structure of the theta method also makes its long-term dynamics amenable to a detailed analysis.

Theorem 1 shows how the linear stability property can be extended to give information about the set of initial conditions and timesteps for which correct, monotonic convergence to steady state is achieved. Theorem 4 concerns spurious period two solutions for small  $\Delta t$  and shows that the kenspeckle value  $\theta = \frac{1}{2}$  forms a precise cut-off between instability and stability. Theorem 5 applies when  $g$  in problem (1) is positive superlinear. In this case the numerical method is simulating a monotonic blow up solution. The theorem shows that period two solutions exist in, and only in, the stable case  $\theta > \frac{1}{2}$ . These stable, oscillatory solutions will clearly lead to a qualitatively incorrect approximation to monotonic blow up, as illustrated in Figure 2. In Theorem 6,  $g$  is taken to be negative superlinear, so that problem (1) is dissipative. The theorem shows that period two solutions exist in, and only in, the unstable  $\theta < \frac{1}{2}$  regime, where dissipativity is then lost. Figure 3 illustrates this phenomenon. These results formalize and extend some of the comments in [10, Section 1] and the examples in [10, Section 2].

Loosely, on blow up problems a stable period two solution ensnares iterates into a spurious oscillatory mode, and on dissipative problems an unstable period two solution repels large initial data away from the correct attractor. In both cases the period two solution is negatively impacting the dynamics.

For the sublinear cases in Theorems 8 and 9, we see that for small  $\Delta t$  period two solutions do not exist, and the theta method behaves well.

Although the results in Section 3 are derived for ODEs, in Section 4 we extended the basic stability cut-off to the case of spatially uniform period two solutions on a general class of semi-discretised PDES.

### *References*

1. M. A. AVES, D. F. GRIFFITHS and D. J. HIGHAM, ‘Runge–Kutta solutions of a hyperbolic conservation law with source term’, *SIAM J. Sci. Stat. Comput.* To appear. 28, 29, 31, 31, 41, 41
2. D. F. GRIFFITHS, P. K. SWEBY and H. C. YEE, ‘On spurious asymptotic numerical solutions of explicit Runge–Kutta methods’, *IMA J. Numer. Anal.* 12 (1992) 319–338. 28
3. E. HAIRER and G. WANNER, *Solving ordinary differential equations II; stiff and differential-algebraic problems*, 2nd edn (Springer Verlag, 1996). 27, 28
4. A. R. HUMPHRIES, ‘Spurious solutions of numerical methods for initial value problems’, *IMA J. Numer. Anal.* 13 (1993) 263–290. 28, 28, 32
5. A. ISERLES, ‘Stability and dynamics of numerical methods for nonlinear ordinary differential equations’, *IMA J. Numer. Anal.* 10 (1990) 1–30. 28, 28, 28
6. J. D. LAMBERT, *Numerical methods for ordinary differential systems* (John Wiley and Sons, Chichester, 1991). 28
7. W. LINIGER, ‘Global accuracy and A-stability of one- and two-step integration formulae for stiff ordinary differential equations’, Proceedings, Conference on the Numerical Solution of Differential Equations, Lecture Notes in Mathematics 109 (ed. J. Ll. Morris, Springer-Verlag, 1969) 188–193. 28, 28
8. J. M. SANZ-SERNA and J. G. VERWER, ‘A study of the recursion  $y^{n+1} = y^n + [y^n]^m$ ’, *J. Math. Anal. Appl.* 116 (1986) 456–464. 36



9. A. M. STUART and A. R. HUMPHRIES, *Dynamical systems and numerical analysis* (Cambridge University Press, Cambridge, 1996). 28
10. A. M. STUART and A. T. PELOW, 'The dynamics of the theta method', *SIAM J. Sci. Stat. Comput.* 12 (1991) 1351–1372. 28, 31, 31, 32, 33, 42, 42

Graeme J. Barclay [ra.gbar@maths.strath.ac.uk](mailto:ra.gbar@maths.strath.ac.uk)

Department of Mathematics  
University of Strathclyde  
Glasgow, G1 1XH

David F. Griffiths [dfg@mcs.dundee.ac.uk](mailto:dfg@mcs.dundee.ac.uk)

Department of Mathematics  
University of Dundee  
Dundee, DD1 4NH  
<http://www.mcs.dundee.ac.uk:8080/dfg/homepage.html>

Desmond J. Higham [djh@maths.strath.ac.uk](mailto:djh@maths.strath.ac.uk)

Department of Mathematics  
University of Strathclyde  
Glasgow, G1 1XH  
<http://www.maths.strath.ac.uk/aas96106/>